Mobility, Mixing and Ergodicity: A Physically-Motivated Measure for Economic Mobility

Viktor Stojkoski [*] Alexander Adamou [†] Yonatan Berman [‡] Colm Connaughton [§]
Ole Peters [¶] [‖]

May 20, 2020

Abstract

Keywords: mobility, inequality, ergodicity economics

[*] Macedonian Academy of Sciences and Arts, vstojkoski@manu.edu.mk
[†] London Mathematical Laboratory, a.adamou@lml.org.uk
[‡] London Mathematical Laboratory, y.berman@lml.org.uk
[§] London Mathematical Laboratory and University of Warwick, c.p.connaughton@warwick.ac.uk
[¶] London Mathematical Laboratory and Santa Fe Institute, o.peters@lml.org.uk
[‖] We thank...

1 Introduction

Economic mobility describes "dynamic aspects of inequality." (?) In the intergenerational context it seeks "to measure the degree to which a child's social and economic opportunities depend on his parents' income or social status." (?) Mobility measures traditionally quantify how wealth (or income[1]) ranks of individuals evolve over time. Intuitively, when mobility is high, ranks evolve quickly, and the chances of an individual to change her position in the wealth distribution over a given time period are high. When mobility is low, individuals are unlikely to change their rank in the distribution over time, or that it changes slowly.

? described several desired properties of statistical measures of mobility and set the standard for such measures. Mobility measures are assumed to be derived from a transition matrix, a bistochastic matrix describing the conditional probabilities of individuals to move between different ranks over some time period. An example for such a measure, used extensively in the mobility literature, is the rank correlation, the correlation between individual wealth ranks in two points in time. Another canonical measure of mobility, used most typically in studies of intergenerational income mobility, is the intergenerational earnings elasticity (IGE), defined as the regression coefficient between log-incomes of parents and children.[2]

The standard statistical measures of mobility have several limitations. First, they are generally incomparable. For example, a rank correlation of 0.3 would have a different meaning if it corresponds to a period of one year or of ten years. Second, for a given time period, it is not generally possible to tell whether the rank correlation is high or low, as it is a dimensionless quantity between -1 and 1. The rank correlation over some time period can only be high or low in comparison to other economies over a similar time period, or to other time periods of the same length.

The interpretation of the rank correlation also depends on the underlying wealth distribution and its dynamics. For example, the same rank correlation cannot be interpreted similarly when the underlying wealth distribution remains unchanged over time, and when it becomes less equal over time.

The IGE has similar limitations, but also others. Most notably, the IGE is sensitive by-design to the level of inequality, *i.e.* to the shape of the underlying distributions, and not only to the transition matrix, as discussed in detail in the mobility literature (*e.g.* **?**).

This paper introduces mixing time, a property of stochastic processes, as a measure of mobility. When wealth is an ergodic observable (?), and assuming that the wealth distribution approaches a steady state, if the wealths of an arbitrary subgroup of individuals are followed over time, the distribution of wealth within this subgroup will converge to the steady-state wealth distribution over time. The convergence time of this process is the mixing time. Put simply, it is the time scale over which individuals mix into the wealth distribution.

[1]We focus on wealth in this paper, but it applies also to income

[2]In fact, the rank correlation and the IGE are both measures of immobility, and to consider them as measures of mobility one has to subtract them from 1.

Using mixing time as a measure for mobility overcomes the comparability issue. If we are interested in mobility over a specific time window – one year, ten years, one generation, *etc.*– determining whether mobility is high or low is immediate by comparing the time window to the observed mixing time. When mixing is rapid, *i.e.* the mixing time is short relative to the window of observation, we interpret mobility as high. Slow mixing is interpreted as low mobility.

We then consider Reallocating Geometric Brownian Motion (RGBM (**???**)) as a model for wealth dynamics and study mixing in this model. In RGBM, individual wealth undergoes random multiplicative growth, modeled as Geometric Brownian Motion (GBM), and is reallocated among individuals by a simple pooling and sharing mechanism. RGBM is a null model of an exponentially growing economy with social structure. It has three parameters representing economic growth, random shocks to individual wealth, and economic interaction among agents, quantified by a reallocation rate. This model is known to reproduce several important stylized facts. In particular, when the reallocation rate is positive, the wealth distribution converges to a stationary distribution with a Pareto tail. The model has both ergodic and non-ergodic regimes, characterized by the sign of the reallocation rate parameter (**?**).

We find that in RGBM the mixing time scales with the inverse of the reallocation rate. As the reallocation rate increases, *i.e.* when a larger share of each individual's wealth is pooled and shared per unit time, the mixing time becomes proportionally shorter, and mobility increases. As the reallocation rate approaches zero, the mixing times get longer, and mobility gets lower. Since decreasing reallocation rates also lead to increasing inequality, this result is in line with the empirical observation that as inequality increases mobility decreases, and vice versa (**?**).

YB: Here we need to describe how mixing time and rank correlation or other standard measures are related in RGBM YB

In practice, many economic observables are best modeled as non-ergodic (**?**). In particular, **?** argue that the US economy is best described in RGBM as one in which wealth is systematically reallocated from poorer to richer, *i.e.* the reallocation rate is negative. In such a case there is no mixing, so the mixing time is infinite. Thus, measuring mobility using standard measures under this regime is misleading. The thorough study of RGBM in this regime is outside of the scope of this paper and left for future work.

The paper is organized as follows. Section 2 discusses the concept of mixing time and how it provides a physically-motivated measure for mobility. Section 4 studies mobility using mixing times in reallocating geometric Brownian motion as a model for wealth. We conclude in Section **??**.

2 Standard measures of economic mobility

2.1 Quantifying economic mobility

Standard measures of economic mobility examine the properties of the bivariate joint distribution describing the wealth of the population in two different periods. Various methods have been developed for evaluating mobility under these circumstances, which, in general, exhibit similar characteristics. Therefore, in order to depict these characteristics we are going to utilize the three most widely used measures of economic mobility: the Spearman rank correlation, the intragenerational earnings elasticity and the wealth transition matrix.

Spearman rank correlation: The Spearman rank correlation between the wealths corresponding to two different time periods, t_m and t_n ($t_m < t_n$) is defined as

$$r_{t_m,t_n} = 1 - \frac{6\sum_i \left(rg(x_i(t_m)) - rg(x_i(t_n))\right)^2}{N(N^2-1)},$$

where $rg(\mathrm{x})$ is the rank transformation of x, $x_i(t)$ is the wealth of individual i in period t and N is the population size. This measure is bounded between -1 and 1, with lower values suggesting greater economic mobility.

Intragenerational earnings elasticity: The second measure is the intragenerational earnings elasticity, quantified via the slope b_{t_m,t_n} of the regression

$$x_i(t_n) = b_0 + b_{t_m,t_n} x_i(t_m) + u_i,$$

where b_0 is the intercept and u_i is the error term. As with the rank correlation, lower intragenerational earnings elasticity also indicates greater mobility. However, this measure is unbounded and may take on any real values.

Wealth transition matrix: The last measure, disaggregates wealth rankings and summarizes economic mobility in a transition matrix $\mathbf{A}$ in which cell entries A_{kl} show the probability that an individual in wealth quantile k in period t_m is found in wealth quantile l in period t_n. In a perfectly mobile economy, the entries of the transition matrix are all equal to each other. In an immobile economy, on the other hand, the largest values are concentrated in the diagonal entries.

2.2 Limitations

To illustrate the limitations of the standard measures, in Fig 1 we construct an artificial example where we examine the economic mobility in a population of 9 individuals during 3 time periods.

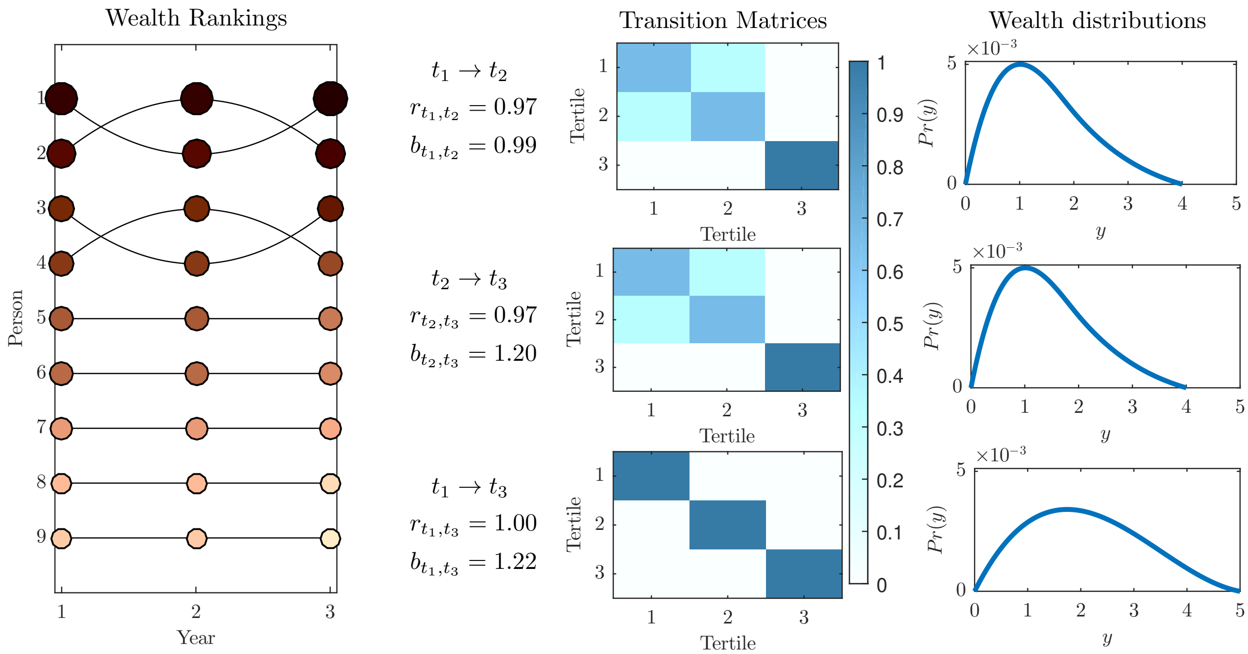

Figure 1: **Performance of mobility measures. Add caption.**

In the left-most panel of the figure, we show the dynamics of the wealth rankings. Besides this, we also highlight the amount of wealth owned by a person with a colored circle, with darker colored and larger circle implying a wealthier person.

In our artificial economy, between the three periods, only the four individuals with largest wealth are able to change their status. In particular, in the second period, t_2 the second wealthiest person becomes the one with the most wealth, while the wealthiest person in the first period t_1 fails to second position. In addition, the third and fourth wealthiest persons change position. In period t_3, the dynamics are reversed, thus ending up with the same wealth rankings as in t_1. The only difference is that in this period, the wealth inequality increased, *i.e.*, the rich got richer and the poor got poorer.

As a result, the rank correlation tells us that the mobility was the same between t_1 and t_2, and t_1 and t_3, but overall between t_1 and t_3 there was no mobility ($r_{t_1,t_3} = 1$). The intragenerational earnings elasticity gives us similar results. In fact, since inequality increased in the last period, it further suggests that the economy became less mobile.

The same conclusions hold when looking at the transition matrices estimated by dividing the wealth rankings in tertiles. An astonishing observation of the last measure is that the transition matrices between t_1 and t_2 and t_2 and t_3 suggest that there is a 1/3 chance for an person belonging in the second tertile to climb to the richer tertile and vice versa. However, we structured our economy in a way that allows movement only between the persons belonging on the edge of the tertile. Hence, the transition matrices fail to adequately represent the movement of the typical person in the tertile.

Obviously, standard measures of economic mobility represent aggregate values of the changes in the wealth rankings of the individuals which constitute the population between two time periods. Therefore, they i) rely on a relevant time period for which the wealth rankings are compared, and ii) do not quantify the mobility of the typical representative of the population.

3 Mixing time

3.1 What is a mixing time?

Mixing time is able to overcome the limitations of standard economic mobility measures by providing an estimate for the characteristic time scale over which individuals mix into the wealth distribution. Think of the economy as a cup of coffee and the typical person's wealth as the milk you pour in the coffee. Mixing time quantifies the time required for the milk to blend with the coffee. This enables the measure to be used for appropriate comparison between different time periods and economies.

As we will see in the following when we will study the RGBM model, there is a direct relation between mixing time and standard mobility measures, whenever mixing time is a finite quantity. However, the standard measures may still indicate that there is mobility even when mixing time does not exist. This is because, mixing time depends on the existence of mobility between *every* quantile in the wealth distribution, whereas for standard measures it is enough for to have transitions only between two quantiles for mobility to exist. In other words, mixing time not only describes the capabilities of poor individuals to rise and become leading social actors, but also evaluates the time required for a rich person to fall in the social ladder. Therefore, we believe that our measure embodies the true concept of mobility.

3.2 Measuring mixing time

In physical terms, mixing describes the property of a dynamical system of being strongly intertwined. That is, any set of particles moving according to the laws of the dynamical system and satisfying a small spread of uncertainty in their initial conditions, follow paths that enter into any region of the phase space, and in a relatively "uniform" way. After a sufficient period of time, the mixing time, the percentage of the particles found within a particular region, is proportional to the volume of that region.

As defined, mixing is strongly related to the concept of ergodicity. However, the latter is a much broader concept: A dynamical system is said to be ergodic if time-averages are equal to ensemble averages. Put differently, ergodicity only implies that every trajectory spreads around the phase space. Hence, every dynamical system that is mixing is also ergodic, but the opposite is not necessarily true.

The ergodicity indicates that in a mixing economy there is always a stationary distribution to which some rescaled transformation of a person's wealth converges. This characteristic can be utilized to develop an estimation procedure for mixing time. Fig 2 gives an example for how to estimate mixing time, whereas the formal procedure is as follows:

1. **Select stationary distribution:** Set a target distribution to which the rescaled transformation of wealth converges (black line in the inset plots);

2. **Select typical representatives:** Select a subsample of the population constituted of people close to the typical representative (red dashed line in the inset plots). For example, this can be the persons whose wealth is closest to either the mean/median/mode wealth, depending on the properties of the stationary distribution;

3. **Track wealth dynamics:** Track the wealth dynamics of each person in the subsample for a sufficient amount of time. Snapshots for the wealth distribution at several time periods are given in the inset plots of Fig 2.

3. **Quantify the differences between the subsample and the stationary wealth distributions:** At certain time periods quantify the differences between the subsample wealth distribution and the stationary distribution. This can be done by utilizing usual statistical distance measures such as the Kullblack-Leibler divergence, the Hellinger distance or the Kolmogorov-Smirnov statistic. In Fig 2, we plot the the log of one such statistic as a function of time with a blue line. In a mixing economy the statistic will exhibit three states. First, there will be a transient state in which the differences between the subsample and stationary wealth distributions will be large and stable. Next, there will be a mixing time state during which the two distributions will slowly converge towards each other, and the statistic will slowly decrease. Finally, there will be a stable state with no significant differences between the stationary and subsample wealth distributions;

1. **Estimate mixing time:** Estimate the slope of the regression in which the dependent variable is the log of the statistic an the independent is time (black dash-doted line in the figure). Take the data only for the mixing time state. Mixing time is the additive inverse of the reciprocal value of the slope ($-1/slope$).

4 Mixing in a simple model of an economy

RGBM is a simple model of a closed one-generation economy. Under RGBM, the dynamics of the wealth $x_i(t)$ of each individual i at time t is specified as

$$\mathrm{d}x_i = x_i\left(\mu \mathrm{d}t + \sigma \mathrm{d}W_i\right) - \tau\left(x_i - \langle x\rangle_N\right)\mathrm{d}t, \tag{4.1}$$

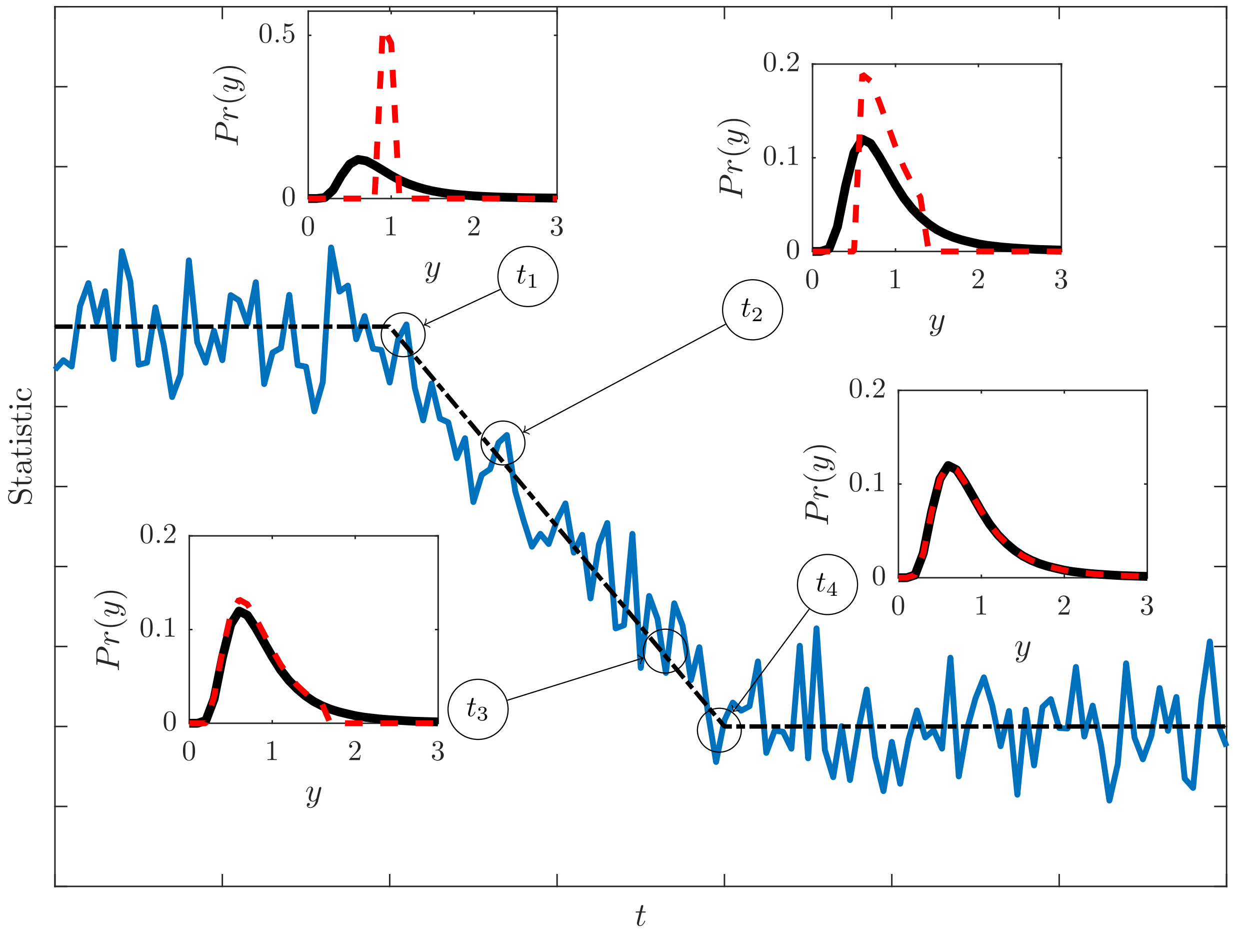

Figure 2: **Mixing time.** Note: y-axis should be log of statistic. Add caption.

with $\mu > 0$ being the drift term, $\sigma > 0$ the noise amplitude, and $\mathrm{d}W_i$ is an independent Wiener increment, $W_i(t) = \int_0^t \mathrm{d}W_i$. In the equation τ is a parameter that quantifies reallocation of wealth. It implies that every year, everyone in the economy contributes a proportion τ of their wealth to a central pool, and then the pool is shared evenly across the population ($\langle x \rangle_N$ is per-capita wealth). The parameter aggregates a multitude of effects: collective investment in infrastructure, education, social programs, taxation, rents paid, private profits made etc.

As can be seen from equation (4.1), the model does not account for a large amount of important characteristics that may constitute an economy, such as openness of the economy (trade with other economies), intergenerationality (interactions between different generations) or directs effects of economic policies. Instead, it focuses solely on the wealth dynamics that are due to individual growth and a consequence of interactions between the individuals. While this may be seen as a drawback of the model, we believe that in fact it is the major advantage of RGBM. It allows us to isolate the effect of non-ergodic wealth dynamics and investigate the resulting implications. Indeed, this characteristic has allowed RGBM to be used in explaining a variety of other social and natural phenomena, ranging from evolution of cooperation **??** up to ontogenetic mass properties **??**.

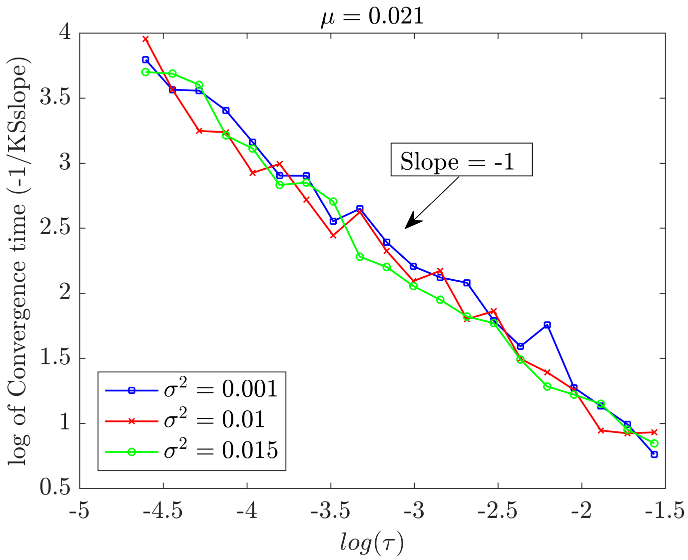

Figure 3: Mixing time in RGBM.

The mathematical properties of RGBM when $\tau \geq 0$ are rather known. In particular, for $\tau > \frac{\sigma^2}{2}$ the model is ergodic and exhibits mean-reversion. Each x_i reverts to the population average $\langle x \rangle_N$. The large population approximation $\langle x(t) \rangle_N = \exp[\mu t]$ is valid, and rescaled wealth $y_i = x_i / \langle x \rangle_N$ has a stationary probability distribution,

$$p(y) = \frac{(\theta-1)^\theta}{\Gamma(\theta)} \exp\big(-\frac{\theta-1}{y}\big) y^{-(1+\theta)}, \tag{4.2}$$

where $\theta = 1 + \frac{2\tau}{\sigma^2}$ and $\Gamma(\cdot)$ is the Gamma function. This distribution exhibits a power law tail and in probability theory is known as the Inverse gamma distribution.

Without reallocation ($\tau = 0$), the model is just GBM. Under GBM wealth is non-ergodic and it follows a lognormal distribution which broadens indefinitely over time. There is no stationary non-zero distribution to which rescaled wealth converges. In GBM inequality is always increasing and mobility is always decreasing.

As pointed out, recent empirical evidence suggests that we are currently living in a negative τ regime **?**. Not much is known about this regime except that there is a self-averaging time period during which individual trajectories repel from the population mean. This introduces negative individual wealth, a phenomenon observed in almost every modern economy, and makes the system non-ergodic. The properties of the model after the self-averaging period are unknown. In what follows, we examine these properties from both an analytical and numerical perspective.

4.1 Mixing time and standard measures of mobility

plot performance of standard measures, and mixing time

4.2 Mobility and inequality in reallocation geometric Brownian motion

The ergodicity, and hence, stationarity allows us to use the stationary distribution in order analytically quantify standard welfare indices and subsequently use them to derive economic policies. For instance, the variance of the distribution can be used as a measure of economic inequality **?**. Its expression is

$$\mathrm{var}(y) = \frac{\sigma^2}{2\tau - \sigma^2}. \tag{4.3}$$

Formally, economic inequality is defined as the extent of concentration in the distribution of wealth among the population. In this aspect, a higher variance implies that the total wealth in the economy is concentrated in few individuals, i.e., the society is more unequal.

In addition, the correlation function of the dynamics can be directly implemented as a measure of intragenerational social mobility, **describe connection between correlation function and measures of economic mobility**

$$\mathrm{corr}(\mathrm{y(t)}, \mathrm{y(t+\delta)}) = \exp\left[-\tau\delta\right], \tag{4.4}$$

where δ quantifies the length of the period used for comparison of the wealth between individuals **?**. Intragenerational social mobility measures the the feasibility of an individual to change their position in the wealth distribution. Since the system is ergodic, the correlation function can be used for deriving the mixing time, i.e., the time scale over which an individuals moves across the whole distribution. This is the relevant measure of immobility in the wealth dynamics of RGBM as it quantifies the amount of time needed for an individual to experience every wealth rank in the population. Its value is $1/\tau$, and a higher value indicates that it takes longer for an individual to visit all possible states of the distribution, thus making the economy less socially mobile.

Fig. 4 confirms the stationary dynamics of these two measures by plotting them as a function of time. The inset plot of the figure gives an example for the analytical power of the positive τ regime in RGBM. Namely, the figure visualizes the stationary Great Gatsby curve. This curve measures the relationship between inequality and social immobility, and has been used for developing social and financial policies. RGBM appropriately reproduces empirical observations by suggesting that economic inequality and social immobility are positively related **?**.

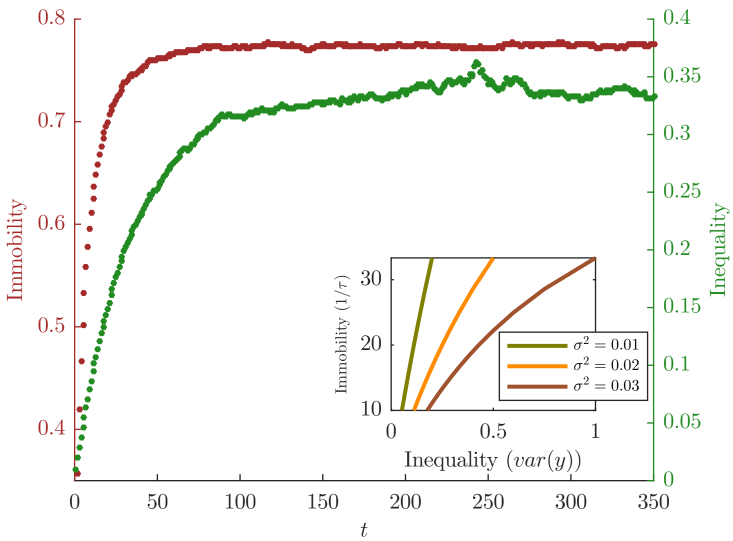

Figure 4: Immobility and inequality in RGBM. **Scatter lines:** Numerical estimation for the inequality and immobility, measured respectively via Eqs. (4.3) and (4.4), in a simulation or RGBM. The inset plot give the analytical relationship between inequality and mobility in RGBM, i.e. the Great Gatsby curve. **Parameters:** We set $\mu = 0.021$, $\sigma^2 = 0.01$ and $N = 10^5$. The initial condition $x_i(0) = 1$ for all i.